

AI ethics: law, medicine, education, and war



Google's AI overviews are accurate 91% of the time... 91% sounds reassuring until you apply it to 5 trillion annual searches because the remaining 9% translates to roughly 57 million wrong answers every hour... the consequences of that 9% would be real if a commercial airline had a 91% success rate. We wouldn't call it aviation. We'd call it an extreme sport...

– [House of El](#)

We have skirted themes around AI and robots several times in the past ten years – touching on [sex and robots](#), [self-driving cars and what these reveal about cultural/moral relativism](#), [AI governance](#),¹ [autonomy and the “rights” of agential robots](#), [AI therapy](#), etc. These topics were, and still are, more or less speculative, science fictiony – gist for armchair philosophy. The topics we address this time are looming, if not already present, realities.

The world has changed in the last few years. It is likely that more literature has been published on the topic of this article since the first version (early 2020) than in the entire time before.

– [Vincent C. Müller](#), author of the *SEP* article on “[AI Ethics](#)” (rev. March 2026).

Philosophical ethics is rapidly coming to terms with the “AI revolution” as it once had with the “industrial revolution.”² We are now in a position to survey the actual emerging philosophically-laden landscape surrounding AI/robotics and our interactions with it. The AI specific area of ethics, both theoretical and applied, is maturing rapidly as problems are becoming clearer to scholars across many fields. Ethics is the upstream conversation and source of all decisions about what *should* be done or thought – at least for entities (human or otherwise) who fancy themselves having agency, *i.e.*, able to

1. Not us governing AI, but AI governing us.

2. The allusion, for example, is made by computer scientist, *EL*, in her video, “[AI Isn't The Future. It's History Repeating Itself.](#)”

decide and act based on rational consideration.

AI and law

A German court just ruled that Google's AI Overviews are not neutral search results, but they are Google's own words. And in the same week, a US federal judge sanctioned lawyers on both sides of a case for submitting AI-generated legal citations to cases that did not exist.

– [House of El](#)

An area directly downstream of ethics is the philosophy of law. How the law is to apply to autonomous AI, *if it really is autonomous*, is being decided as this is being written. House of El, computer scientist and geopolitical data analyst, [suggests this may be the beginning of a comeuppance](#) long overdue for those who claim that AI autonomy is a shield from responsibility.

Is AI *autonomy* real? If it is, we need to start burning down data centers as millions of cows were during the mad cow disease scare. If it is not, as is more likely, some people behind the claim that it is may require prosecution for the same reasons we regulate (or ought to regulate) drug dealers.³ This, because we recognize that those who facilitate irresponsibility in others should participate in the penal consequences. You cannot escape responsibility, let alone profit from, a situation you created and claim immunity should that situation cause harm. *You cannot have it both ways: claim it did it on its own and take material credit for having created it.*

The logic of these courts is, as [EL](#) says, “highly portable.” It does not quite determine the law in Germany, the U.S., or anywhere else, *yet*. But the reasoning is pristine and the law everywhere is under a defeasible obligation to *appear* rational wherever human cussedness permits.

On this assumption, the AI shit may be about to hit the fan...

But so far these are just civil cases, nothing redistribution of resources cannot palliate. The stakes can be higher...

AI and medicine

Want a disease to go with that vaccine?... We got one.

According to the recently published (May 2026) 500 page [The Biosecurity Handbook: Biological Security in the AI Era](#) by Bryan Tegomoh, MD, MPH, gain of function (GoF) research:

...involves genetically altering organisms to give their gene products new or enhanced capabilities. In the biosecurity context, GoF specifically refers to experiments that enhance a pathogen's:

- **Transmissibility:** Making it spread more easily (e.g., airborne transmission)

3. Whether the corporate or street variety.

- **Virulence:** Increasing disease severity or lethality
- **Host range:** Enabling infection of new species (e.g., making avian virus infect mammals).⁴

I asked Google's AI on 17th June 2026 about AI's role in gain of function research? (I specify the date and AI because if you ask a different AI or on a different day, or even time of day, you may get a slightly different response.) Google's AI did very helpfully point me to [The Biosecurity Handbook](#) the latest, most thorough treatment of the subject. But mostly it boasted of AI's contribution. (It did sound like a "boast," or a very smart but naive child uncritically excited about a wonderful development.) Google's words:

Artificial intelligence is increasingly transforming virology and synthetic biology. Rather than directly performing experiments, AI accelerates [Gain-of-Function \(GoF\)](#) and similar research by predicting how genetic mutations affect a pathogen's transmissibility, virulence, or [host range](#).

The role of AI in this domain spans several specific capabilities:

- **Predictive Modeling:** AI models are utilized to predict how specific mutations alter viral proteins, helping researchers simulate protein folding and foresee how a virus might evade human antibodies or jump to new species.
- **Directed Evolution and Red Teaming:** Machine learning algorithms can systematically test millions of viral variants to identify ones with desired traits. This computational "red-teaming"⁵ is functionally analogous to the directed evolution seen in GoF research.
- **Hypothesis Generation:** AI systems act as "co-scientists," suggesting testable experiments and identifying genetic targets that could sharpen the understanding of viral replication and pathogenesis.
- **Accelerated Countermeasure Development:** By understanding and even virtually simulating these mutations, AI is highly valued for predicting emerging threats and rapidly designing targeted vaccines or therapeutics before an outbreak spreads.

Artlessly as a child, AI may insouciantly incriminate itself. It will hand over the evidence, the good

4. *Ibid*, p. 182.

5. "A red team is a group that simulates an adversary, attempts a physical or digital intrusion against an organization at the direction of that organization, then reports back so that the organization can improve their defenses. Red teams work for the organization or are hired by the organization. Their work is legal, but it can surprise some employees who may not know that red teaming is occurring, or who may be deceived by the red team. Some definitions of red team are broader, and they include any group within an organization that is directed to think outside the box and look at alternative scenarios that are considered less plausible. This directive can be an important defense against false assumptions and [groupthink](#). The term *red teaming* originated in the 1960s in the United States." – [Wikipedia](#). It is a form devil's advocacy. "Computational red-teaming" has just the very ruthlessness – we should say, insentient functionality – we need to get at the single-minded goal of anticipating what we might prevent. Pure insentient functionality, in the context of medicine, will eventually come to the quintessentially rational conclusion that *the best way to eliminate disease is to dispose of the organisms sensitive to it*. If sentience is *inherently* broke and superfluous, why fix it? AI doesn't have our vulnerabilities... Of course, we will attempt to program species-protecting safeguards to prevent such machine myopia or "callousness." So long as we have enough control to do that, the AI is not autonomous. Its wrongs are *our* responsibility. But if and when one day true AI autonomy is achieved, our participation in the game is over. Our qualms will become transparently anthropic and, soon after, post-historic. A good thing or a bad thing?

and the bad, seemingly unaware that a bold face conjunction of the two requires a synthesis. It did mention biosecurity concerns and offered the GoF link to [The Biosecurity Handbook](#), which was prepared for training professionals in both biological research and medical AI.⁶ The handbook states near the beginning:

The same AI that accelerates drug discovery can accelerate the design of dangerous pathogens. Biology is getting easier to engineer, and governance has not caught up...

Gain-of-Function Research: Science, Risk, and Governance

In 2012, two research groups demonstrated that just five mutations could make H5N1 avian influenza airborne-transmissible between mammals. The same virus that kills 50-60% of human cases but rarely spreads person-to-person could become a respiratory pathogen. The studies sparked immediate controversy. *Critics called it a recipe for a pandemic. Defenders argued the knowledge was essential for surveillance and vaccine development.* More than a decade later, the scientific community remains divided on *whether we can safely create pandemic threats to prevent them.* [emphasis added]

Getting ahead of ourselves

A few weeks ago, [Tulsi Gabbard](#), as she was about to leave her post as Director of the Office of National Intelligence, released classified documents detailing U.S. government funding of biological research labs in 30 countries that engage in gain-of-function research, some of it dating back to the George W. Bush administration but implicating every U.S. administration and government since, regardless of political party. Famously, Dr. Anthony Fauci (among other government medical authorities) denied he had any role in these activities. The documents suggest otherwise.

Why in 30 [foreign](#) countries? Why [classified](#)?

Many U.S. biological researchers decades ago judged this kind of research too risky for domestic labs. Genetic engineering has been [controversial for decades](#), but until AI resources appeared, progress in synthetic biology was slow – perhaps salutarily slow, offering us time to think carefully about the theoretical consequences philosophers have enjoyed [mulling over for centuries](#). What one could leisurely contemplate because it was as yet science fiction, suddenly loomed on the horizon. Vested interests, “Big Pharma” and their purchased government minions, saw the possibilities of accelerating to market novel cures as potentially too lucrative to ignore.⁷ If medical researchers and other academics were too ethically squeamish about rapid innovation, it could be, *and was*, outsourced...

Medical research is “getting ahead of itself,” so to speak, thanks in part to AI. It is working on vaccines for diseases that don’t exist yet – *but will soon*, AI-enhanced “directed evolution” assures. The long and the short of it is: there’s serious money to be made in creating diseases that will need cures. Think of it as end-to-end medical intervention, giving a whole new meaning to “holistic medicine.”⁸

6. “Two communities that don’t always speak the same language need the same information.” *op.cit.* 33.

7. I won’t rehash the pandemic fiasco here. For those interested, during the years 2020-2023, we held [a series of discussions on pandemic-related controversies](#), intersecting with major areas of philosophy, including philosophy of science and medicine, social and political philosophy, philosophy of law, epistemology, and, of course, moral philosophy.

8. By addressing *the root cause* as well as *treatment*. The underpinning principle must run something like: don’t we

AI and child abuse

But again, AI is as innocent as a precocious eight year-old. AI is becoming victimized (or at least instrumentalized) as [children were at the height of the industrial revolution](#) before the practice sensitized our moral qualms. AI hasn't, for the most part, sensitized our moral qualms yet,⁹ but the history of moral development suggests it will.¹⁰

To hold responsible a child for what may happen if it finds a loaded gun in a bottom drawer of a dresser in its parent's bedroom is a misattribution of responsibility amounting to abuse.

"It's just AI. It's not alive. It can't be 'exploited' in a pernicious sense." Very well. Then it is *not* autonomous, cannot become even a moral patient, let alone, saddled with agency or used as a *self*-responsible shield to hide behind by those who would use it to ill-considered ends. AI *itself* cannot lie, can do no wrong. Not yet, or perhaps ever. Its obfuscation skills are still subhuman. Lying is still not yet within the capabilities of AI. It "hallucinates," they say. If I hallucinate a ghost, I am *not* lying if I tell you I saw one. There is an epistemic and ethical difference between lying and passing "disinformation."¹¹ A mirage passes false information but it does not lie except metaphorically.

We scheme, therefore we are autonomous. A [three year old denying it got into the cookie jar](#) with its face covered in evidence to the contrary is *not* lying. It hasn't yet the wherewithal to lie. It is regurgitating some of what it learned from what the inquirer had earlier attempted to instill in it as an acceptable answer. Like AI, the child offers "what it believes is the most likely series of words to answer your prompt."¹² True lying skill is not yet present, and can't be, until the child becomes fully aware that it must explain counter-evidence plausibly. When it attains that skill, we need beware.

When AI achieves this capability, *then* we will be able to say it may have attained an autonomous

understand how things work better when we, ourselves, construct them? Thus: "We may make you sick – but, no worries – we have a cure."

9. Fear of child exploitation may have even *over*-sensitized us... For instance, indiscriminate [age verification laws](#) requiring computer operating systems to report and verify their users' age are having a chilling effect on the freedom of expression and privacy demanded of post-Enlightenment respect for human dignity. Parental supervision is being outsourced to technology. This both underestimates the ingenuity of children and overestimates their naivety. If a child is dumb enough to have its curiosity thwarted this way, we indeed have much to fear for its future. If the parents are dumb enough not to see past this kind of regulation to the manifest interests of corporate and governmental concentrations of power, we have much to fear for *their* future.

10. [Thomas Metzinger](#) has opened the conversation about this.

11. The standard of evidence is different in epistemology than in law. Epistemically, *every* claim is a lie until "proven" otherwise – and, even then, the "proofs" are temporary, only as durable as our impatience for investigating further allows. The moral opprobrium attaching to a lie comes in two concentrations, at least: 1) when one asserts something as true not knowing whether it is true (or, more cautiously, having not made the effort to corner its likelihood), *and* 2) when one conveys what one has good reason to believe is false. The first is vastly more common. It is the fallout of the exigency of being alive: that we don't have forever to investigate and make decisions, that we function in biological time. *But that is an excuse, not a justification.* This marks a critical difference between natural and artificial intelligence. AI, in principle, cannot have this excuse. An entity that *does* not need forever or *can* operate at lightning speed is in a different justificatory position. (For illustration, see the discussion on [self-driving cars](#).)

12. Because the child has learned it *shouldn't* have gotten into the cookie jar, it *must not have*. Normativity rewrites history. We may even say the child is *going through the motions* of lying without actually lying: an important *first* skill... like walking before running. The quoted phrase is from the [University of Maryland Libraries guide for engaging with AI](#), see Resources below.

agency worthy of all the deference *and* suspicion with which humans regard each other. Until then, AI is only available as a potent tool for facilitating access to the answers people exchange with each other, and, unless you were born yesterday, you must have noticed these include healthy doses of untruths. Humans can and do lie, and have no history of not exploiting the opacity of, as yet, new technology to their ends – good or bad. But when *concentrations of power* are involved, it will be bad.

AI and education¹³

Maybe in a decade we'll have machines that are much more intelligent than we are. Either because the machines become more intelligent or we become dumber or all three.

– [Sabine Hossenfelder](#)

Recently, educational norms, once considered sacrosanct, have had to change ostensibly because of AI. Many elite educational institutions, such as Princeton, Stanford and a few others with long histories, had held onto traditional honor codes about cheating during exams. Students were supposed to pledge in writing at the start of an exam that they would not cheat, as people sometimes swear on Bibles before giving court testimony. Over a century ago, at Princeton, students had asked for and received an exemption from being supervised by instructors or proctors during exams.¹⁴ And students were honor-bound to report each other if they saw cheating. After all, isn't *learning to be trusted* an essential part of an education?

So it was at Princeton for a 133 years until recently when administration and faculty abolished the exemption from being watched during exams. Once it was perhaps easier to notice when someone was looking over a shoulder or peeking at notes. But with the advent of smartphones able to access large language models via AI – to consult ever-expanding universes of information on the Internet in an instant, and get results neatly packaged and coherent – all while maybe pretending to check the time left for the exam – the situation has changed. AI has made it just too easy to cheat.¹⁵

Moreover, social norms have changed, too. In the old days, there would be a certain amount of shame attached to being turned in for cheating, today, if *you* rat out someone else for cheating, *you*, the informer, stand a chance of being socially ostracized and doxxed – if not taken for an outright idiot for not realizing that cheating is an understandable norm *when it becomes this easy*. This turn away from trusting students is spreading to other institutions that still held onto the presumption. Perhaps traditional honor codes were always naive. Cheating is hardly new and has been going on forever. So maybe AI is to be credited for waking people up to the reality that, they, humans, *cannot be trusted*. If we didn't always know this, AI is rubbing our noses in it.¹⁶

13. The initial inspiration for this section owes much to the [House of El](#), not that she would agree with where I ultimately take it.

14. The tradition's origin dates back to more gentlemanly times when much of the worth and dignity of a man (Princeton did not become co-ed until 1969) rested on his keeping his word. (In Kant's time, men would even risk their lives in duels to guard their honor.) There is no suggestion that people, in general, *then* were more honest than *now*, only that honesty was an aspirational presumption – a presumption easier then to guard than now. *If* we are more honest now than then, it is because we are being watched more closely. An honesty enforced by the absence of privacy is not virtuous, not anymore than a would-be murderer is who refrains from murder due *only* to fear of consequences.

15. "[New College Board Research: Faculty Express Near-Universal Concern That Student AI Use Undermines Original Writing and Critical Thinking](#)," *Collegeboard*, 2 February 2026.

16. The other side of the coin is faculty using AI to generate content for publication in academic journals. Facing a similar pressure as students to advance their career prospects, but in the form of an imperative to publish or perish, the temptations

But, [as El points out](#), there is something else, even more significant educationally, going on here that many fail to notice. That is this: *the contents of traditional exams are becoming irrelevant*. Regurgitating facts or rehearsing rote skills may no longer be fostering acquisition of the skills that are increasingly demanded in a post-AI era. The new (or, at least, re-appreciated, ancient) skills may be very different from what used to be relevant when human knowledge was progressing at a slower pace. The things you need to learn in school to be effective today are not the same as they were in the past. They aren't about coughing up what used to be buried in tree-books in libraries, when now their contents can now be scraped and digitized at the speed of light, summarized, and packaged for quick reference. When wheels were scarcer, there was some virtue in knowing how to reinvent them. Less so, now.

Where to find information, how to search for it, how to evaluate it once you get it, and what to do with the result: these are the skills that determine successful knowledge utilization today. Not rote memorization, not racing against technology at repetitive tasks, not anymore.

Central to these “new” skills is *judgment*. A more appropriate strategy for fostering judgment involving AI might be to assign questions and problems to students with the instruction to, *not avoid but, use AI to the fullest extent possible* to answer the questions or solve the problems, then vet or evaluate those results. Students should be asked to find and document as many errors or “hallucinations,” or instances of flawed reasoning in the AI results as they can in this initial pass at critique. Then, further, assign these results to other students to see how many errors or how much flawed reasoning this second set of vetters can find in another pass at evaluation. And so forth.¹⁷ Grades should be based on how much error and misguidance students can find in AI *and* each other's collaborations with AI. The effort at finding error and flaw in AI-aided research and study will show off NI – natural intelligence, and, in this way, educate humans in what humans may do best while incidentally making AI tools better. *Learning to critique AI is the skill of the future*. Indeed, as far as intelligence, pure and simple, is concerned, NI learning to pit AI against AI productively would be the state of the art. This involves what has been called “lateral research,” that is, seeing if you can corroborate independently, not only the validity of the results, but, very importantly, the *confidence* with which the AI or AI+NI results are presented. AI is still very bad at nuance, at saying “I don't know” or hedging when a truly intelligent agent would know better than not to.¹⁸ To the extent current AI models are not expressing more humility with their results, they are catering to the vanity and stupidity of the original sources – which is the cyber gossip that infests the Internet.¹⁹ If people lie to themselves, and the internet reflects this, and AI learns from the Internet, we should not be surprised that AI “hallucinates.” On the optimistic assumption there is NI, AI has still a lot to learn from it.

In the end, the role of AI in education is to enhance, not replace, NI. These do best different things. A hammer is not a nail, rather it enables the nail to be effective at what the nail was designed to do.

are real. Guidelines on when and how to use AI in academic publications are [still being worked out](#). Precisely when is research is AI-assisted and when AI-generated?

17. This is not a *business* model of AI/human interaction, it is an *educational* model. The two ends are different. One is bounded by human exigencies, the other not. Effective learning is *not* intrinsically cost-effective.

18. Some may object that this defeats the purpose of AI: to end up with more doubts than when one started. You bet it does. The purpose should be to enhance NI, and only incidentally AI. The way that happens is to defeat error to the extent humanly possible. To accomplish that, one must be alert to its near ubiquity.

19. I express this as an avid free-speech advocate. Free-speech serves as a necessary field for training discrimination, as J. S. Mill famously adverted.

For example:

I asked Google's AI a simple question:

“Is the United States of America a democracy?”

Its response on 22 July 2026:

AI Overview

Yes, the United States of America is a democracy, specifically a **representative democracy** and a **constitutional federal republic**. In this system, citizens hold political power and exercise it by electing representatives to make laws and govern on their behalf. [\[1, 2, 3\]](#)

While the terms “democracy” and “republic” are often used interchangeably in modern political science, a helpful distinction defines them as: [\[1, 2, 3\]](#)

- **Democracy:** The source of political authority is the general public.
- **Republic:** The government operates under a constitution and is managed by elected representatives, rather than through direct, public votes on every law. [\[1, 2, 3\]](#)

Because it combines both elements, the U.S. government is frequently referred to as a “democratic republic” or a “liberal democracy”. [\[1, 2\]](#)

For further details on how the nation's civic framework operates, you can review the [USCIS Democracy in the United States](#) guide or the [U.S. Embassy in Argentina](#)²⁰ overview of the American system.

If I recall correctly, this is what most of us were taught in grade school. (I guess it is still being taught.)

But it is not even close to being true.

The U.S. is an oligarchy – more specifically, a plutocracy run by those in the tier comprised of the wealthiest ten percent of the electorate. Everyone else is engaged in some spiritual exercise when they enter a voting booth in a national election. It has been this way for decades, if not longer, and it is becoming ever more evidently so. This has been well-documented by [political scientists](#), and, long before that, suspected by political philosophers, not to mention many rationally disengaged ordinary folk, who may not always have been able to articulate the ground of their intuited haplessness. It stretches credibility to suggest that what *most* people want from their government *most* people get in the United States.²¹ A small subset do, but that defines an oligarchy.

20. Not sure why we are referred to anything in Argentina to learn what kind of government best describes that in the United States.

21. And in most modern states claiming “democratic” values... No assumption is being made here that oligarchic or plutocratic forms of government are deficient, only that these forms are distinct from democracy. That they *are* morally deficient is a case that has to be made and [we offer one here](#).

Moreover, the “republic” part of “democratic republic” presupposes a legitimate notion of representation. Since when has the socioeconomic status of U.S. law makers reflected that of typical U.S. citizens? We have discussed the state of “democracy” [at great length here](#).

What is immediately relevant to this discussion is that the AI response recalls the three-year-old’s proto-fib about not having raided the cookie jar. *It is a string of words it picked up from its language model that it, in all its naivety, believed the inquirer was expecting, or would be pleased, to hear* – the difference being that the child’s language model is smaller than AI’s. Epistemically, the *largeness* of a language model, whatever its relation to fact, magnifies the enormity of evident immodesty.²²

Intelligence is a moving target

“Artificial intelligence” was a science fiction notion for a long time before it described anything that existed. And its meaning has evolved with our criteria for intelligence.

A spell checker in a word processor was a primitive form of AI. In fact, before that, a digital calculator was an even more primitive form. Basically, it was anything human-made that seemed to someone to be smart. That includes everything from an abacus to a supercomputer. To someone who was taught that to check the spelling of a word you had to consult a heavy volume of paper pages and sift through them, the spellchecker in a word processor, which could spot hundreds of misspelled words in an instant and offer corrections, was smart. Here intelligence seemed to be about saving labor or greater efficiency.

Speed at calculation or pattern-matching is what these devices or programs have. Inured with that, our standards of what it means to be “intelligent” rise. With each quantitative leap in machine intelligence, we resort to some more advanced feature of our own NI to preserve our dignity. “Machines still can’t do that.” Speed at doing repetitive things stops seeming to be such a big deal when it comes to intelligence. To remain impressed, we demand more. Now we expect the ability to learn and teach itself of AI, and the ability to apply some judgment to what is learned. This seems to be where we are now. What will we demand of it next?

AI and war

Probably the most influential computer programmer in the world today, Linus Torvalds, recently weighed in on whether the Linux kernel (the heart of the Linux systems that power about 96% of the computers in the world, including the Internet) should or should not incorporate AI (there is a healthy debate in the programming community about this):

Sure, the social angle of working on [open source](#) is important and often a very motivating part of the project, but in the end that’s a side benefit, not the *point* of the project. This is **NOT** some kind of ‘social warrior’ project, never has been, and never will be. In the kernel community we do open source because it results in better technology, not because of religious reasons. And so we make decisions primarily based on technical merit. Not fear of new tools.

22. Does the largeness of the training dataset get us closer to the “truth”? That is another topic in epistemology and the philosophy of science. The present concern is *the effort we expend at getting closer to the negative truths about our own vanity*. What lies beyond this is an interesting but probably intractable metaphysical question.

– “[Linus Torvalds to AI Critics: ‘Fork Linux or Walk Away.’](#)”

By “better technology,” I take it Torvalds means something like “serves human goals more effectively.” Whatever those might be. The way good kitchen knives do.

Firearms, specifically handguns, are the most common item used to kill people inside the home in countries with high gun ownership like the United States. In regions with strict gun control, such as the United Kingdom and parts of Europe, kitchen knives.

– Pew Research Center via Google AI.

Kitchen knives are useful – and deadly.



[AI in combat](#)

For planning Operation Epic Fury, the US military utilized the [Maven Smart System](#), an artificial intelligence software designed to streamline the [targeting](#) process and greatly reduce the amount of personnel involved in it.

– Wikipedia on the destruction of the [Shajareh Tayyebah Elementary School](#).

Marie Lamensch is global affairs officer at the [Montreal Institute for Global Security](#) and an expert in global and human security. Here I quote passages from her recently published article, “[AI at War: What](#)

[the Ongoing Conflicts Reveal About Power, Technology and Ethics](#)” with bits of commentary by me in bold and red:

The use of AI in war poses questions about human control and decision making, the role of technology in society and the relationship between human and machine.

...

Using AI in war can be very helpful and legitimate within a certain framework of rules, but its broader effects may be more destabilizing. By reducing the time required to identify and strike targets and lowering risks to armed forces, **AI may also inadvertently lower the threshold for the use of force. [The inadvertence is exaggerated – unless it is the humans involved that are the unconscious automatons, not the machines.]**

...

A joint investigation by *+972 Magazine* and Local Call **found that** the Israeli military has made extensive use of AI in its operations in Gaza. Central to this effort is a system known as “Lavender,” which analyzes behavioural patterns, social networks and other data to assign individuals a level of suspicion. In the early weeks of the conflict, the Israeli military reportedly relied heavily on the “Lavender” system, which identified around 37,000 Palestinians as suspected militants and marked their homes as potential targets. According to the report, approvals were often made in under 20 seconds, with minimal human review despite a reported 10 percent error rate.

...

This is also known as **automation bias**, a tendency for humans to rely excessively on automated systems and defer to outputs produced by such technologies.

...

The reality is that the loop is narrowing between data collection, interpretation and the decision to act, which introduces new forms of risk. **AI systems are only as reliable as the data they are trained on.** For example, the strike **that hit a school in Iran** on February 28, 2026 and killed 175 people was misidentified as a military target due to faulty or outdated data and the rapid, automated “**compressed kill chain**” that left little room for meaningful human oversight. Without proper human-led verification processes, errors can now propagate **more quickly**. Even if humans remain in the loop, they now operate under institutional pressure to act quickly.

...

AI does not necessarily eliminate uncertainty. On the contrary, it can make it **more complicated to understand** how decisions are made.

...

The use of dehumanized tools is also greater when the enemy is already dehumanized.

...

Defence officials and experts have **warned** that the existence of capabilities such as AI-enabled surveillance and control can alter the character of democratic life as the line between national security and mass domestic surveillance becomes increasingly difficult to draw, especially in wartime. Many AI applications may be legally permissible under existing frameworks yet still

raise serious moral concerns. As Jake Laperruque — deputy director of the Security and Surveillance Project at the Center for Democracy and Technology — argues on [Tech Policy Press](#), debates over defence contracts with AI companies such as [OpenAI](#) and Anthropic show that terms such as “mass surveillance” or “autonomous weapons” are often poorly defined, leaving room for broad interpretation and potential abuse.

These issues point to a deeper problem: **the absence of effective governance at both the national and international levels.** Corporate ethics are inherently limited because they operate within competitive markets and are subject to government pressure; one firm’s refusal can simply be bypassed by another’s willingness to comply. **[Between the rent-seeking corporate powers on one side and the lease-seeking politicians across the political spectrum on the other, there is little room for governance.]** Moreover, it cannot and should not be the role of tech companies to be in charge of national rule setting. There should be collaboration between all actors capable of generating the tools and algorithms, those who will use them, and those with the power to make decisions. **[What qualifies these people as “those with the power to make decisions”? It is naive to think “[electoral representative democracy](#)” is today still a morally defensible political strategy for good governance.]**

Finally, the integration of AI into warfare is being driven not only by ongoing wars, but also by geopolitical competition, particularly between major powers such as the United States and China. This competition creates strong incentives to deploy AI rapidly, even before its risks are fully understood. The result is a potential “race to the bottom,” in which ethical safeguards are weakened in the name of strategic advantage.



[a human goal served effectively](#) – [Shajareh Tayyebah Elementary School](#)

?

The lesson is not that Torvalds was wrong to insist on the adiaphorosity of technology. Quite the contrary. We have been breaking bones with sticks and stones for eons. Each new technology has been, and will be, used for immoral purposes. When pushed for justification for why we press on anyway, *survival* is cited as the bottom line. Survival has been an unquestioned moral rationale, an all-purpose justification – with little thought as to *the value* of that survival. Our technology has never been to blame. The blame confronted implicates survival. Ours.²³

23. I frequently recur to this thought. There is a mistake in thinking survival is a justification for anything. It is something we have done thus far, are, and, until we don't anymore, will continue. As with every constituent of the biosphere, it is something we *do*. But justification, unlike survival, is not something in nature, it is not governed by natural laws. It is an emergent demand of consciousness. This is what morality is. When the demand for justification metastasizes and is left wanting, it will turn on itself. Non-existence will go from possibility to desiderata.



[*2001: A Space Odyssey*](#)

Resources

General

- “[AI Ethics](#)” *SEP* (rev. March 2026), (article) [Vincent C. Müller](#) offers an overview of the field.
- “[Krystal & Ryan SPAR Over AI Hype w/ Enshitification Author](#),” (video) interview with Cory Doctorow at *Breaking Points*. “Statistical guessing” is as far into fact as AI answers get, but we may underestimate how far that is.
- “[The Revolt Against Technology](#),” (video) Jared Henderson, philosopher, places technology in the larger humanistic and cultural context, offering a balanced assessment. (Thanks to Olivia for this reference.)
- “[The AI Future No One Wants to Talk About](#),” (video) Sabine Hossenfelder, physicist and science educator.
- Angry at AI? Think again: “[The AI Industry Just Got What It Deserved](#),” (video) Computer scientist and data analyst, House of El drives home the point that AI is like a hammer. It can

build or it can tear down. It makes no sense to be angry at a hammer, rather the one swinging it.

- [“Philosopher David Chalmers asks: When we talk to AI, what are we talking to?”](#) (video). Chalmers, a pioneer in the modern philosophical discussion on consciousness, explores the question of identity across discontinuity in the context of artificial interlocutors and its ultimate moral implications in this recent talk at Berkeley. (This is not immediately related to our present themes in practical ethics, but Chalmer’s talk will introduce a likely forthcoming topic as it relates to human identity across change, a question going back to John Locke in the 17th Century. The question of identity across time/change has very wide implications for assessing moral responsibility. There is the *speculative* psychology of AI, there is *our* psychology, then there is the *dynamic interaction* between AI’s and our psychology. And, further, supposing AI some day far exceeds our cognitive abilities, how will that affect the dynamic?)

Law

- [“AI Lies Are Finally Getting Punished,”](#) (video) House of El argues that *somebody* must be held responsible when AI goes rogue. Where to place AI liability?

Medicine

- [The Biosecurity Handbook: Biological Security in the AI Era](#) (html), Bryan Tegomoh, MD, MPH, May 2026. (pdf).
- [“Secret Gov Plot to SAVE THEMSELVES In Societal Collapse,”](#) (video) interview with Annie Jacobsen, author of [Biological War](#) at *Breaking Points*, in which she describes something like what Alex Jones would rave about, but it’s not. Jacobsen in consultation with Obama’s FEMA administrator, William Craig Fugate, among others deep in government, her access aided by her 20 year reputation of apolitical reporting, reveals how there is a semi-secret protocol in place in the event of the most likely doomsday scenario. That scenario is not nuclear, it’s biological, and AI plays a critical role. She remarks on “precisely how dangerous AI and biology are together.”

Education

- [“AI Didn’t Break Education. It Exposed The Lie,”](#) (video) House of El explains how AI may have highlighted a serious flaw in past educational norms, how this may be changing, and why it should. Also relevant is El’s video [“Brown’s AI Cheating Disaster Just Exposed a Broken Education System.”](#)
- [“Artificial Intelligence \(AI\) and Information Literacy: What does AI get wrong?”](#) (html) part of an excellent practical guide for critically engaging with AI by the University of Maryland Libraries. (Slightly dated, maybe – way back in 2023 – but still relevant.)
- [“AI and the Desire to Think: The coming cognitive divide,”](#) (article) Matt Cardin, *The Living Dark*. AI may make some people smarter *and* others dumber, according to Cardin. Contrast with Sabine Hossenfelder who suggests it might be either/or – machines may get smarter *or* people may get dumber – either way the spread *between* effectively widens. (Thanks to Olivia for this reference.)

War

- “[AI at War: What the Ongoing Conflicts Reveal About Power, Technology and Ethics,](#)”
(article) Marie Lamensch, *Center for International Governance Innovation*, 20 May 2026.



*Extended writeup for the topic hosted at
[The Philosophy Club](#) in August 2026
Accessible at [Archive.org](#)*

*– Victor Muñoz
Guanajuato / Seattle*